Small deviations of the determinants of random matrices with Gaussian entries[1]

Nadezhda V. Volodko[2]

Abstract

The probability of small deviations of the determinant of the matrix AA^T is estimated, where A is an $n \times \infty$ random matrix with centered entries having joint Gaussian distribution. The inequality obtained is exact in a sense.

1 Introduction and the result.

We discuss the problem of estimating probabilities of small deviations for the determinants of random matrices of a special type. Regarding the above-mentioned topic, there are papers devoted to the small deviations of stationary Gaussian processes with respect to different norms (see [2], [3] and the references therein) or to the analogous problem for the smallest singular values of random matrices (see [1] and the references therein).

Theorem. *Let $A = \{\tau_{ij}\}$ be an $n \times \infty$ random matrix with centered entries having joint Gaussian distribution. Let*

$$\mathbf{P}\Big(\sum_{j=1}^{\infty} \tau_{ij}^2 < \infty\Big) = 1$$

for every $i \le n$. Suppose also that τ_{kk} cannot be represented as a linear combination of the entries τ_{ij}, with $\min(i,j) < k$, i.e., for each $k \le n$,

$$d_k := \inf_{\{\alpha_{ij}\}} \mathbb{E}\Big(\tau_{kk} - \sum_{\min(i,j)<k} \alpha_{ij}\tau_{ij}\Big)^2 > 0.$$

Then, for any $\varepsilon > 0$,

$$\mathbf{P}(\sqrt{\det AA^T} < \varepsilon) \le C_1^n \varepsilon_0 \Big(\sum_{k=0}^{n-1} \frac{|\log(C_2^n \varepsilon_0)|^k}{k!}\Big), \tag{1}$$

[1]This work was supported by RFBR grants N 11-01-00285, 11-01-12006-ofi-m-2011.

[2]Sobolev Institute of Mathematics of the Russian Academy of Sciences. E-mail: nvolodko@gmail.com

where C_1 and C_2 are absolute positive constants,

$$\varepsilon_0 = \frac{\varepsilon}{\prod_{i=1}^n |d_i^{1/2}|}.$$

Remark. *In the case of diagonal $n \times n$-matrix with independent Gaussian entries, the following lower estimate is valid:*

$$\mathbf{P}(\sqrt{\det AA^T} < \varepsilon) \geq c^n \varepsilon_0 \Big(\sum_{k=0}^{n-1} \frac{|\log \varepsilon_0|^k}{k!} \Big),$$

where c is an absolute constant.

2 Proof.

Proof of Theorem. At first, prove the following simple fact: adding to any (infinite) row of the matrix A a linear combination of other rows does not change the determinant $\det(AA^T)$. Denote by A_k the k-th row of matrix A. Then $AA^T = \{\langle A_i, A_j \rangle\}$. Further, if we replace the last row A_n by $A_n + \sum_{k=1}^{n-1} \alpha_k A_k$ then only the last row and the last column of the matrix AA^T will be changed (due to the symmetry, they coincide after transposition). Consider the elements of the last row for $i \leq n-1$:

$$\langle A_i, A_n + \sum_{k=1}^{n-1} \alpha_k A_k \rangle = \langle A_i, A_n \rangle + \sum_{k=1}^{n-1} \alpha_k \langle A_i, A_k \rangle;$$

$$\langle A_n + \sum_{k=1}^{n-1} \alpha_k A_k, A_n + \sum_{k=1}^{n-1} \alpha_k A_k \rangle = \langle A_n, A_n \rangle + 2 \sum_{k=1}^{n-1} \alpha_k \langle A_n, A_k \rangle + \sum_{i,k=1}^{n-1} \alpha_i \alpha_k \langle A_i, A_k \rangle.$$

Thus, the last row and the last column of the transformed matrix AA^T are obtained from the last row and the last column of the initial matrix by adding the linear combination of the previous $n-1$ rows (or columns, correspondingly). Then the determinant does not change after these transformations.

Now, execute the Gram-Schmidt orthogonalization of the set of the infinite-dimensional vectors $\{\tau_{ij}\}_j$. Let $\{B_i(j)\}_j$ be such that

$$\{B_1(j)\}_j \equiv \{\tau_{1j}\}_j,$$

and for $i \geq 2$ and, for all j,

$$\tau_{ij} = \alpha_1^i B_1(j) + ... + \alpha_{i-1}^i B_{i-1}(j) + B_i(j) \text{ a. s.},$$

and moreover, for $i \neq j$,

$$\langle B_i, B_j\rangle := \sum_{k=1}^{\infty} B_i(k)B_j(k) = 0 \text{ a. s.} \tag{2}$$

It is easy to note that, for $i \geq 2$,

$$B_i(k) = \tau_{ik} - \frac{\langle A_i, B_1\rangle}{\langle B_1, B_1\rangle} B_1(k) - \ldots - \frac{\langle A_i, B_{i-1}\rangle}{\langle B_{i-1}, B_{i-1}\rangle} B_{i-1}(k) \text{ a. s.} \tag{3}$$

Due to the argumentation above,

$$\det(AA^T) = \langle B_1, B_1\rangle \cdot \ldots \cdot \langle B_n, B_n\rangle.$$

Denote the determinant under consideration by $\det_n$ and prove the assertion of Theorem by induction. For $n = 1$ the assertion is valid, now let it be fulfilled for $\det_{n-1}$ with $n \geq 2$. We need to prove (1) for $\det_n$. Due to the Theorem conditions there exists a standard normal variable ζ independent from τ_{ij} with $\min(i, j) < n$ (and consequently, from $\det_{n-1}$), such that for $k \geq n$

$$\tau_{nk} = a_n^k \zeta + \overline{\zeta}_k,$$

where $\overline{\zeta}_k$ and ζ are independent, $(a_n^n)^2 = d_n > 0$.

$$\mathbf{P}(\sqrt{\det_n} \leq \varepsilon) = \mathbf{P}\Big(\langle B_n, B_n\rangle \leq \frac{\varepsilon^2}{\det_{n-1}}\Big) = \mathbf{P}\Big(\sum_{k=1}^{\infty} B_n^2(k) \leq \frac{\varepsilon^2}{\det_{n-1}}\Big). \tag{4}$$

For the convenience of notation, put $a_n^k = 0$ for $k \leq n-1$, and $a_n = \{a_n^k\}_{k\geq 1}$. Then, due to representation (3), we can write $B_n^2(k)$ as

$$B_n^2(k) = \Big(\Big(a_n^k - \sum_{i=1}^{n-1} \frac{B_i(k)}{\langle B_i, B_i\rangle} \sum_{j=n}^{\infty} B_i(j)a_n^j\Big)\zeta + \rho_k\Big)^2, \tag{5}$$

where ρ_k and ζ are independent. Prove the following assertion:

$$\mathbf{P}\Big(\sum_{k=1}^{\infty}\Big(a_n^k - \sum_{i=1}^{n-1} \frac{B_i(k)}{\langle B_i, B_i\rangle} \sum_{j=n}^{\infty} B_i(j)a_n^j\Big)^2 \leq \frac{u^2}{\det_{n-1}}\Big)$$

$$= \mathbf{P}(d\sqrt{\det_{n-1}} \leq u) \leq C_1^n u_0 \Big(\sum_{k=0}^{n-2} \frac{|\log(C_2^n u_0)|^k}{k!}\Big) \tag{6}$$

where

$$d = \Big(\sum_{k=1}^{\infty} \Big(a_n^k - \sum_{i=1}^{n-1} \frac{B_i(k)}{\langle B_i, B_i \rangle} \sum_{j=n}^{\infty} B_i(j) a_n^j \Big)^2 \Big)^{1/2},$$

$$u_0 = \frac{u}{\prod_{i=1}^{n} |a_i^i|}.$$

For $u_0 = 1$ suppose that $0^0 = 1$.

$$\mathbf{P}(d\sqrt{\det_{n-1}} \le u) \le \mathbf{P}\Big(\sum_{k=1}^{n} \Big(a_n^k - \sum_{i=1}^{n-1} \frac{B_i(k)}{\langle B_i, B_i \rangle} \sum_{j=n}^{\infty} B_i(j) a_n^j \Big)^2 \le \frac{u^2}{\det_{n-1}} \Big)$$

$$\le \mathbf{P}\Big(\Big\{ \sum_{k=1}^{n-1} \Big(\sum_{i=1}^{n-1} \frac{B_i(k)}{\langle B_i, B_i \rangle} \sum_{j=n}^{\infty} B_i(j) a_n^j \Big)^2 \le \frac{u^2}{\det_{n-1}} \Big\}$$

$$\cap \Big\{ |a_n^n - \sum_{i=1}^{n-1} \frac{B_i(n)}{\sqrt{\langle B_i, B_i \rangle}} \frac{\langle B_i, a_n \rangle}{\sqrt{\langle B_i, B_i \rangle}}| \le \frac{u}{\sqrt{\det_{n-1}}} \Big\} \Big). \tag{7}$$

Introduce the notations:

$$B^0 := \Big\{ \frac{B_i(j)}{\sqrt{\langle B_i, B_i \rangle}} \Big\}_{i,j \le n-1}, \quad y := \Big\{ \frac{\langle B_1, a_n \rangle}{\sqrt{\langle B_1, B_1 \rangle}}, ..., \frac{\langle B_{n-1}, a_n \rangle}{\sqrt{\langle B_{n-1}, B_{n-1} \rangle}} \Big\}^T.$$

Obviously,

$$\Big| \frac{B_i(k)}{\sqrt{\langle B_i, B_i \rangle}} \Big| \le 1,$$

and then

$$\langle B^0(k), B^0(k) \rangle = \|B^0(k)\|^2 \le n - 1.$$

Here $B^0(k)$ is the k-th row of matrix B^0. Also, we will denote

$$B^0(n) = \Big\{ \frac{B_i(n)}{\sqrt{\langle B_i, B_i \rangle}} \Big\}_{i \le n-1}$$

Then, denoting the event under the probabilbity sign on the right hand side of (2) by Ω_0, write the following estimate:

$$\mathbf{P}(\Omega_0) \le \mathbf{P}(\Omega_0; \sqrt{\det_{n-1}} > 2u/|a_n^n|) + \mathbf{P}(\sqrt{\det_{n-1}} \le 2u/|a_n^n|). \tag{8}$$

Due to the induction assumption for inequality (1) the value $\mathbf{P}(\sqrt{\det_{n-1}} \le 2u/|a_n^n|)$ does not exceed the right hand side of (6). Consequently, for the proof of (6) it is sufficient to estimate only the first summand of (8).

Fix an elementary outcome from the intersection of events Ω_0 and $\{\sqrt{\det_{n-1}} > 2u/|a_n^n|\}$. Event Ω_0 means the fulfillment of the system of two inequalities. Analyse the system mentioned on the chosen elementary outcome. Due to the conditions $a_n^n \neq 0$ and $\sqrt{\det_{n-1}} > 2u/|a_n^n|$, the last inequality of the system means that

$$|B^0(n)y| \geq |a_n^n|/2,$$

and consequently,

$$\|y\| \geq \frac{|a_n^n|}{2\|B^0(n)\|}.$$

The first inequality of the system can be briefly written as follows:

$$\|v\| := \|B^0 y\| \leq \frac{u}{\sqrt{\det_{n-1}}}.$$

Note, that due to the Theorem conditions $\mathbf{P}(\det B^0 = 0) = 0$. Then

$$\frac{a_n^n}{2\|B^0(n)\|} \leq \|y\| = \|(B^0)^{-1}v\| \leq \max_{i \leq n-1} |\lambda_i|^{-1/2}\| \leq \max_{i \leq n-1} |\lambda_i|^{-1/2} \frac{u}{\sqrt{\det_{n-1}}},$$

where λ_i are eigenvalues of symmetric matrix $(B^0)^T B^0$. It is clear that

$$\min_{i \leq n-1} |\lambda_i| \leq \frac{4\|B^0(n)\|^2 u^2}{(a_n^n)^2 \det_{n-1}}. \tag{9}$$

Due to the property of the matrix trace and the representation

$$(B^0)^T B^0 = \Big\{ \frac{\sum_{k=1}^{n-1} B_i(k) B_j(k)}{\sqrt{\langle B_i, B_i\rangle \langle B_j, B_j\rangle}} \Big\},$$

the following assertion is valid:

$$\sum_{i \leq n-1} \lambda_i \leq n - 1 - \|B^0(n)\|^2. \tag{10}$$

Prove the next easy assertion: if

$$y_i \geq 0,\ i = 1, ..., k; \quad \sum_{i=1}^{k} y_i = A$$

then

$$\prod_{i=1}^{k} y_i \leq \Big(\frac{A}{k}\Big)^k. \tag{11}$$

Prove (11) by induction: for $k = 1$ it is evident, now let it hold for $k - 1$. Then

$$\prod_{i=1}^{k} y_i = \prod_{i=1}^{k-1} y_i \cdot y_k \leq \Big(\frac{A - y_k}{k-1}\Big)^{k-1} y_k.$$

The product vanishes at the points $y_k = 0$ and $y_k = A$. Then, at the point inside the interval $[0, A]$, where the derivative by y_k of the expression above equals zero, the expression mentioned reaches its maximum:

$$\Big(\Big(\frac{A - y_k}{k-1}\Big)^{k-1} y_k\Big)' = \Big(\frac{A - y_k}{k-1}\Big)^{k-1} - (k-1)y_k\frac{(A-y_k)^{k-2}}{(k-1)^{k-1}} = \Big(\frac{A - y_k}{k-1}\Big)^{k-2}\Big(\frac{A - ky_k}{k-1}\Big).$$

We obtain that the point of maximum is $y_k = A/k$, whence (11) follows.

The eigenvalues of $(B^0)^T B^0$ are nonnegative, then, due to (10) and (11), we obtain that

$$\prod_{i=2}^{n-1} \lambda_i \leq \Big(\frac{n - 1 - \|B^0(n)\|^2}{n-2}\Big)^{n-2}.$$

Here without loss of generality we suppose $\lambda_1 = \min_{i \leq n-1} |\lambda_i|$. Consequently,

$$\det(B^0)^T B^0 \leq \frac{4\|B^0(n)\|^2 u^2}{(a_n^n)^2 \det_{n-1}}\Big(\frac{n - 1 - \|B^0(n)\|^2}{n-2}\Big)^{n-2}.$$

Find maximum by $\|B^0(n)\|^2$ of the expression on the right hand side. Similarly note, that the expression vanishes on the edges of permissible interval $[0, n-1]$, then we need to find zero of the derivative inside this interval. After calculations we obtain that maximum is reached at the point $\|B^0(n)\|^2 = 1$. Thus,

$$\det(B^0)^T B^0 \leq \frac{4u^2}{(a_n^n)^2 \det_{n-1}},$$

and obviously

$$|\det B^0| \leq \frac{2u}{|a_n^n|\sqrt{\det_{n-1}}}.$$

If we multiply a matrix row or column by any number, then the determinant is multiplied by this number. Adding to any row (or column) of matrix a linear combination of other rows (columns) does not change the determinant.

Taking into account these well-known facts and representation (3), we obtain

$$\det B^0 = \frac{\det\{B_i(j)\}_{i,j \leq n-1}}{\sqrt{\langle B_1, B_1\rangle} \ldots \sqrt{\langle B_{n-1}, B_{n-1}\rangle}} = \frac{\det\{\tau_{ij}\}_{i,j \leq n-1}}{\sqrt{\det_{n-1}}}.$$

Then,

$$|\det\{\tau_{ij}\}_{i,j\le n-1}| \le \frac{2u}{|a_n^n|}.$$

As we realized calculations on the fixed elementary outcome from the intersection of the events Ω_0 and $\sqrt{\det_{n-1}} > 2u/|a_n^n|$, one can conclude that the first summand on the right hand side of (8) does not exceed the probability

$$\mathbf{P}\Big(|\det\{\tau_{ij}\}_{i,j\le n-1}| \le \frac{2u}{|a_n^n|}\Big). \tag{12}$$

Note that we can apply the induction assumption to the probability above. If one put $f_{ijk} = 0$ with $\max(i,j,k) > n-1$ in the problem statement then τ_{ij} with $\max(i,j) > n-1$ equal to the identical zero and $\sqrt{\det_{n-1}} = |\det\{\tau_{ij}\}_{i,j\le n-1}|$. Moreover, the values d_n can only increase in comparing with the initial matrix. Then, the value in (12) does not exceed the right hand side of (6). Hence it follows that estimate (6) is valid.

Return to the proof of (1), and particularly, to formula (4). Denote the event under the probability sign on the right hand side of (4) as W:

$$\mathbf{P}(W) \le \mathbf{P}(d\sqrt{\det_{n-1}} \le \varepsilon) + \mathbf{P}(W; d\sqrt{\det_{n-1}} > \varepsilon). \tag{13}$$

Estimate the second summand. Denote by P_* a conditional probability given all random variables from (5) for all k except (and independent from) ζ.

$$\mathbf{P}(W; d\sqrt{\det_{n-1}} > \varepsilon) = \mathbb{E}(\mathbf{P}_*(W); d\sqrt{\det_{n-1}} > \varepsilon)$$

$$= \mathbb{E}\Big[\mathbf{P}_*\Big(\sum_{k=1}^{\infty}\Big(\Big(a_n^k - \sum_{i=1}^{n-1}\frac{B_i(k)}{\langle B_i, B_i\rangle}\sum_{j=n}^{\infty}B_i(j)a_n^j\Big)\zeta + \rho_k\Big)^2 \le \frac{\varepsilon^2}{\det_{n-1}}\Big); d\sqrt{\det_{n-1}} > \varepsilon\Big]$$

$$\le \mathbb{E}\Big[\frac{C\varepsilon}{d\sqrt{\det_{n-1}}}; d\sqrt{\det_{n-1}} > \varepsilon\Big]$$

$$\le C\varepsilon\Big(\mathbb{E}\big[(d\sqrt{\det_{n-1}})^{-1}; \varepsilon < d\sqrt{\det_{n-1}} < C_2^{-n}|a_1^1\cdot\ldots\cdot a_n^n|\big] + C_2^n|a_1^1\cdot\ldots\cdot a_n^n|^{-1}\Big).$$

Here and further we will use notation C for different positive absolute constants, whereas C_2 is chosen the same as in (6). Denote the distribution function of $d\sqrt{\det_{n-1}}$ as $F_0(t)$ and estimate the expectation on the right hand side:

$$\mathbb{E}\big[(d\sqrt{\det_{n-1}})^{-1}; \varepsilon < d\sqrt{\det_{n-1}} < C_2^{-n}|a_1^1\cdot\ldots\cdot a_n^n|\big] = \int_{\varepsilon}^{C_2^{-n}|a_1^1\cdot\ldots\cdot a_n^n|}\frac{1}{t}dF_0(t)$$

$$= \frac{F_0(t)}{t}\Big|_{\varepsilon}^{C_2^{-n}|a_1^1\cdot\ldots\cdot a_n^n|} + \int_{\varepsilon}^{C_2^{-n}|a_1^1\cdot\ldots\cdot a_n^n|} \frac{F_0(t)dt}{t^2}$$

$$\leq \frac{C^n}{\prod_{i=1}^n |a_i^i|} - \frac{F_0(\varepsilon)}{\varepsilon} + \frac{C^n}{|a_1^1\cdot\ldots\cdot a_n^n|}\int_{C_2^n\varepsilon|a_1^1\cdot\ldots\cdot a_n^n|^{-1}}^{1} \sum_{k=0}^{n-2} \frac{|\log u|^k}{k!u} du$$

$$\leq \frac{C^n}{\prod_{i=1}^n |a_i^i|} - \frac{F_0(\varepsilon)}{\varepsilon} + \frac{C^n}{|a_1^1\cdot\ldots\cdot a_n^n|}\Big(\sum_{k=1}^{n-1} \frac{|\log(C_2^n\varepsilon_0)|^k}{k!}\Big).$$

Thus, taking into account (6) and the arguments above, estimate (13):

$$\mathbf{P}(W) \leq CF_0(\varepsilon) + C^n\varepsilon_0\Big(\sum_{k=1}^{n-1} \frac{|\log(C_2^n\varepsilon_0)|^k}{k!}\Big).$$

Assertion (1) is proved.

Proof of Remark. In conditions of Remark

$$\sqrt{|\det AA^T|} = |\tau_1\cdot\ldots\cdot\tau_n|,$$

where $\tau_1, ..., \tau_n$ are independent Gaussian variables with dispersions $d_1, ..., d_n$. Obviously, one can represent

$$\tau_k = \sqrt{d_k}\eta_k,$$

where $\eta_1, ..., \eta_n$ are independent standard normal variables. Then

$$\mathbf{P}(|\tau_1\cdot\ldots\cdot\tau_n| < \varepsilon) = \mathbf{P}\Big(|\eta_1\cdot\ldots\cdot\eta_n| < \frac{\varepsilon}{(\prod_{i=1}^n d_i)^{1/2}}\Big),$$

therefore it is sufficient to prove the Remark assertion for the case $d_1 = ... = d_n = 1$ only.

Prove by induction. For $n = 1$ the assertion holds, now let it be valid for $n-1$, where $n \geq 2$. Denote the distribution function and the density of the standard normal law by $\Phi(\cdot)$ and $\varphi(\cdot)$ correspondingly, the distribution function of $|\eta_1\cdot\ldots\cdot\eta_{n-1}|$ by $F_0(\cdot)$, and different positive absolute constants by c.

$$\mathbf{P}(|\eta_1\cdot\ldots\cdot\eta_n| < \varepsilon) \geq \mathbf{P}\Big(|\eta_n| \leq 1; |\eta_1\cdot\ldots\cdot\eta_{n-1}| < \varepsilon\Big)$$

$$+\mathbf{P}\Big(|\eta_n| \leq \frac{\varepsilon}{|\eta_1\cdot\ldots\cdot\eta_{n-1}|}; \varepsilon < |\eta_1\cdot\ldots\cdot\eta_{n-1}| < 1\Big)$$

$$\geq [2\Phi(1)-1]F_0(\varepsilon) + \varphi(1)\varepsilon\mathbb{E}\Big((|\eta_1\cdot\ldots\cdot\eta_{n-1}|)^{-1}; \varepsilon < |\eta_1\cdot\ldots\cdot\eta_{n-1}| < 1\Big)$$

$$\geq [2\Phi(1)-1]F_0(\varepsilon)+\varphi(1)\varepsilon\Big(\frac{F_0(u)}{u}\Big|_\varepsilon^1+c^n\int_\varepsilon^1\sum_{k=0}^{n-2}\frac{|\log u|^k du}{uk!}\Big)$$

$$=[2\Phi(1)-1-\varphi(1)]F_0(\varepsilon)+c^n\varepsilon\sum_{k=0}^{n-1}\frac{|\log\varepsilon|^k}{k!}.$$

Due to the induction assumption and the inequality $2\Phi(1)-1-\varphi(1)>0$, Remark is proved.

References

[1] R. Adamczak, O. Guedon, A. E. Litvak, A. Pajor, and N. Tomczak-Jaegermann, "Condition number of a square matrix with i.i.d. columns drawn from a convex body," Proc. Amer. Math. Soc. **140**, 987–998 (2012).

[2] F. Aurzada, I. A. Ibragimov, M. A. Lifshits, and H. van Zanten, "Small deviations of smooth stationary Gaussian processes," Theor. Probab. Appl. **53**, 697–707 (2008) [www.arXiv.org/abs/0803.4238].

[3] Ya. Yu. Nikitin and R. S. Pusev, "Exact l_2-small deviation asymptotics for some Brownian functionals," Teor. Veroyatn. Primen. **57** (1), 98–123 (2012) [http://arxiv.org/abs/1104.2891].